文章编号:1671-9352(2025)03-0088-12 DOI:10.6040/j.issn.1671-9352.0.2023.226

线性混合效应模型的复合分位数回归估计

李京[1],杨宜平[1,2*],赵培信[1,2]
(1.重庆工商大学数学与统计学院, 重庆 400067; 2.经济社会应用统计重庆市重点实验室, 重庆 400067)

摘要:考虑线性混合效应模型的稳健估计问题,通过结合矩阵的 QR 分解技术和复合分位数回归方法,提出一种基于正交投影的复合分位数回归估计方法。先通过 QR 分解技术消除随机效应,再构造固定效应的复合分位数回归目标函数,从而获得固定效应的估计。在一些正则条件下,证明所提出估计的渐近正态性。所提出的估计方法无需对模型误差和随机效应的分布作任何限制性的假定,并且固定效应的估计不受随机效应的影响。与正交矩估计方法的模拟研究比较表明,本文提出的方法具有稳健性,并将其应用于实际数据分析。

关键词:线性混合效应模型;QR 分解;复合分位数回归;固定效应;随机效应
中图分类号:O212.7 **文献标志码**:A
引用格式:李京,杨宜平,赵培信. 线性混合效应模型的复合分位数回归估计[J]. 山东大学学报(理学版),2025,60(3):88-99.

Composite quantile regression estimation of linear mixed effects model

LI Jing[1], YANG Yiping[1,2*], ZHAO Peixin[1,2]
(1. College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China;
2. Chongqing Key Laboratory of Social Economic and Applied Statistics, Chongqing 400067, China)

Abstract: Considering the robust estimation problem of linear mixed effect model, a composite quantile regression estimation method based on orthogonal projection is proposed by combining the QR decomposition technique of matrix and the composite quantile regression method. The random effects are eliminated by QR decomposition technique, and then the fixed effects are estimated by constructing the composite quantile regression objective function. Under some regular conditions, the asymptotic normality of the proposed estimates is proved. The proposed estimation method does not need to make any restrictive assumptions about the distribution of model errors and random effects, and the estimates of fixed effects are not affected by random effects. Further, the simulation study compares the proposed method with the orthogonality-based estimation of moment method, which shows that the proposed method is robust and applied to the actual data analysis.

Key words: linear mixed effect model; QR decomposition; composite quantile regression; fixed effect; random effect

0 引言

线性混合效应模型广泛地应用于相关数据的分析,比如纵向数据和重复测量数据。随着相关数据研究的不断发展,线性混合效应模型在生物医学、计量经济学、社会科学等各个领域都有着广泛的应用。大量学者关注并研究了线性混合效应模型:Cui 等[1]基于矩估计方法研究了线性混合效应模型中固定效应的估计问题;Wu 等[2]基于 QR 分解针对线性混合效应模型的随机效应以及模型误差的高阶矩提出了一

收稿日期:2023-05-20; **网络出版时间**:2023-11-30 08:53:12
基金项目:国家社会科学基金资助项目(18BTJ035); 重庆市自然科学基金资助项目(cstc2021jcyj-msxmX0079,cstc2020jcyj-msxmX0006); 重庆市教委人文社科一般项目(21SIGH118); 重庆市社科规划委托项目(2019WT58); 第五批重庆市高等学校优秀人才支持计划(68021900601)
第一作者:李京(1999—),男,硕士研究生,研究方向为数理统计. E-mail:1728630092@qq.com
*通信作者:杨宜平(1981—),女,教授,博士,研究方向为非参数统计及数据分析. E-mail:yeepingyang@foxmail.com

种正交矩估计方法;陈心洁等[3]利用 FIC 选择准则方法讨论了线性混合效应模型的变量选择问题;林鹏[4]基于修正 Cholesky 分解的硬阈值估计和一种罚估计考虑了线性混合效应模型中随机效应的选择问题;赵培信等[5]基于矩阵的 QR 分解技术,对含有不完全观测数据的线性混合效应模型提出了一种基于正交投影的估计方法。

目前,大多数关于线性混合效应模型的研究都集中在最小二乘估计的框架下进行讨论的,然而最小二乘估计容易受到离群点的影响,难以满足关于模型误差的假设条件。在实际收集到的数据中,通常有异常值、尖峰、厚尾等异质性,此时采用最小二乘估计的方法将不再具有优良性,且稳定性很差。为了获得稳健性的估计,Zou 等[6]提出了复合分位数回归,该方法是一个有效的稳健性估计方法。复合分位数回归提出后被广泛应用:王康宁等[7]利用 copula 函数和复合分位数回归讨论了纵向数据模型的稳健估计和变量选择;刘艳霞等[8]基于复合分位数回归考虑了部分线性变系数模型中未知参数和非参数分量的估计问题;张永霞等[9]将贝叶斯和复合分位数回归方法相结合,研究了部分线性单指标模型的估计问题;张立文等[10]基于复合分位数研究了门限自回归模型的变点问题;Jiang 等[11]研究了超高维单指标复合分位数回归模型;Guo 等[12]将复合分位数回归推广到了超高维半参数模型平均的研究。

本文研究的问题是针对混合效应模型结合复合分位数回归获得一个稳健的估计,本文提出方法的创新点在于:(1)借鉴 Wu 等[2]的思想,采用 QR 分解技术消除了随机效应,使得固定效应的估计不受随机效应的影响。(2)与 Wu 等[2]提出的估计方法相比,本文构造的复合分位数回归目标函数使得固定效应的估计更加稳健。

1　方法与主要结果

1.1　复合分位数回归估计

考虑如下线性混合效应模型

$$Y_{ij}=\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}+\boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{b}_i+\varepsilon_{ij},\quad i=1,2,\cdots,n,\quad j=1,2,\cdots,l_i, \tag{1}$$

其中:Y_{ij}是响应变量;$\boldsymbol{X}_{ij}$和 $\boldsymbol{Z}_{ij}$分别为 $p\times1$、$q\times1$ 型协变量;l_i 为第 i 个个体的观测次数;$\boldsymbol{\beta}$ 为 $p\times1$ 型固定效应;$\boldsymbol{b}_i=(b_{i1},b_{i2},\cdots,b_{iq})^{\mathrm{T}}$ 为第 i 个个体的 $q\times1$ 型随机效应;ε_{ij}为模型误差,假定$\{\varepsilon_{ij},\ i=1,2,\cdots,n,\ j=1,2,\cdots,l_i\}$和$\{\boldsymbol{b}_i,\ i=1,2,\cdots,n\}$是独立同分布的随机变量序列。

记 $\boldsymbol{Y}_i=(Y_{i1},Y_{i2},\cdots,Y_{il_i})^{\mathrm{T}}$ 是 $l_i\times1$ 型向量,$\boldsymbol{X}_i=(\boldsymbol{X}_{i1},\boldsymbol{X}_{i2},\cdots,\boldsymbol{X}_{il_i})^{\mathrm{T}}$ 是 $l_i\times p$ 型矩阵,$\boldsymbol{Z}_i=(\boldsymbol{Z}_{i1},\boldsymbol{Z}_{i2},\cdots,\boldsymbol{Z}_{il_i})^{\mathrm{T}}$ 是 $l_i\times q$ 型矩阵及 $\boldsymbol{\varepsilon}_i=(\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{il_i})^{\mathrm{T}}$ 是 $l_i\times1$ 型向量,那么模型(1)可写为

$$\boldsymbol{Y}_i=\boldsymbol{X}_i\boldsymbol{\beta}+\boldsymbol{Z}_i\boldsymbol{b}_i+\boldsymbol{\varepsilon}_i,\quad i=1,2,\cdots,n\ 。 \tag{2}$$

进一步地,假设 $\boldsymbol{Z}_i$ 是列满秩的矩阵,下面对矩阵 $\boldsymbol{Z}_i$ 进行 QR 分解,即

$$\boldsymbol{Z}_i=\boldsymbol{Q}_i\begin{pmatrix}\boldsymbol{R}_i\\ \boldsymbol{0}\end{pmatrix},$$

其中:$\boldsymbol{Q}_i$ 是 $l_i\times l_i$ 型正交矩阵;$\boldsymbol{R}_i$ 是 $q\times q$ 型上三角矩阵;$\boldsymbol{0}$ 是$(l_i-q)\times q$ 型零矩阵。现将 $\boldsymbol{Q}_i$ 划分为 $\boldsymbol{Q}_i=(\boldsymbol{Q}_{i1},\boldsymbol{Q}_{i2})$,其中 $\boldsymbol{Q}_{i1}$是 $l_i\times q$ 型矩阵,$\boldsymbol{Q}_{i2}$是 $l_i\times(l_i-q)$型矩阵。由矩阵的正交性可知 $\boldsymbol{Z}_i=\boldsymbol{Q}_{i1}\boldsymbol{R}_i$ 和 $\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{Q}_{i1}=\boldsymbol{0}$,因此有 $\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{Z}_i=\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{Q}_{i1}\boldsymbol{R}_i=\boldsymbol{0}$。基于此,在模型(2)两边同时左乘 $\boldsymbol{Q}_{i2}^{\mathrm{T}}$可得模型

$$\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{Y}_i=\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{X}_i\boldsymbol{\beta}+\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{\varepsilon}_i,\quad i=1,2,\cdots,n。 \tag{3}$$

令 $\tilde{\boldsymbol{Y}}_i=\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{Y}_i$ 是$(l_i-q)\times1$ 型向量,$\tilde{\boldsymbol{X}}_i=\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{X}_i$ 是$(l_i-q)\times p$ 型矩阵,$\tilde{\boldsymbol{\varepsilon}}_i=\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{\varepsilon}_i$ 是$(l_i-q)\times1$ 型向量,则模型(3)可变为

$$\tilde{\boldsymbol{Y}}_i=\tilde{\boldsymbol{X}}_i\boldsymbol{\beta}+\tilde{\boldsymbol{\varepsilon}}_i,\quad i=1,2,\cdots,n。 \tag{4}$$

式(4)通过 QR 分解消除了随机效应,即模型(4)变为仅含有固定效应的线性回归模型。注意到,通过变换后,样本量由 nl_i 变成了 $n(l_i-q)$。为了简单起见,记 $N=n(l_i-q)$。基于模型(4),可以构造 $\boldsymbol{\beta}$ 如下复合分位数回归估计 $\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}$,即

$$(\hat{a}_1,\hat{a}_2,\cdots,\hat{a}_K,\hat{\boldsymbol{\beta}}^{\mathrm{CQR}})=\underset{a_1,a_2,\cdots,a_K,\boldsymbol{\beta}}{\arg\min}\sum_{k=1}^{K}\left\{\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\rho_{\tau_k}(\tilde{Y}_{ij}-\tilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\beta}-a_k)\right\},$$

其中:$\rho_{\tau_k}(r)$是分位数损失函数且$\rho_{\tau_k}(r)=\tau_k r-rI(r<0)$, $k=1,2,\cdots,K$, $0<\tau_1<\tau_2<\cdots<\tau_K<1$; a_k 是 $\tilde{\varepsilon}_{ij}$的 τ_k 分位点;$\tilde{Y}_{ij}$是 $\tilde{\boldsymbol{Y}}_i$ 的第 j 行;$\tilde{\boldsymbol{X}}_{ij}$是 $\tilde{\boldsymbol{X}}_i$ 的第 j 行;K 是分位点的个数。上述目标函数是对分位数回归模型的改进,综合了多处分位数回归的信息,所以估计更稳健。通常,使用等间隔的分位点,即 $\tau_k=\dfrac{k}{K+1}$, $k=1,2,\cdots,K$,等间隔分位点操作方便。除此之外,它包括了一些常见分位点的信息,如 $K=5,9$ 时,包含了常用的分位点如中位数、10%分位点、90%分位点等。Zou 等[6]指出,当 $K\geqslant 5$ 时,估计值接近真实值。综上,本文在模拟研究中取 $K=5,9$ 这 2 种情况进行验证。对于该非线性目标函数的最值问题,可以直接调用 R 软件中的 cqrReg 包来解决。

1.2 渐近性质

为了得到$\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}$的渐近性质,需要以下条件:

(C1) 存在一个 $p\times p$ 型正定矩阵 $\boldsymbol{\Sigma}$,使得

$$\boldsymbol{\Sigma}=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{Q}_{i2}\boldsymbol{Q}_{i2}^{\mathrm{T}}\boldsymbol{X}_i);$$

(C2) 对于任意的 p 维向量 $\boldsymbol{u}$ 和常数 μ, $\tilde{\boldsymbol{\varepsilon}}$ 的累积分布函数 $F(\cdot)$和密度函数 $f(\cdot)$满足

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\int_0^{\mu_0+\tilde{\boldsymbol{X}}_i^{\mathrm{T}}\boldsymbol{u}}\sqrt{n}\left[F\left(a+\frac{t}{\sqrt{n}}\right)-F(a)\right]\mathrm{d}t=\frac{1}{2}f(a)(\mu_0,\boldsymbol{u}^{\mathrm{T}})\begin{pmatrix}1&0\\0&\boldsymbol{\Sigma}\end{pmatrix}(\mu_0,\boldsymbol{u}^{\mathrm{T}})^{\mathrm{T}}。$$

定理 1 若正则条件(C1)、(C2)成立,则有

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}-\boldsymbol{\beta})\xrightarrow{d}N(0,\boldsymbol{\Sigma}_{\mathrm{CQR}}),$$

$$\boldsymbol{\Sigma}_{\mathrm{CQR}}=\boldsymbol{\Sigma}^{-1}\frac{\sum_{k,k'}^{K}\min(\tau_k,\tau_{k'})(1-\max(\tau_k,\tau_{k'}))}{\left(\sum_{k=1}^{K}f(a_k)\right)^2},$$

其中$\xrightarrow{d}$表示依分布收敛。

证明 记$\boldsymbol{\beta}$、a_k 是参数真值,令$\sqrt{n}(\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}-\boldsymbol{\beta})=\boldsymbol{\mu}_n$, $\sqrt{n}(\hat{a}_k-a_k)=v_{n,k}$,则$(v_{n,1},v_{n,2},\cdots,v_{n,K},\boldsymbol{\mu}_n^{\mathrm{T}})$可以通过最小化如下目标函数可得

$$L_N=\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\left(\rho_{\tau_k}\left(\tilde{\varepsilon}_{ij}-a_k-\frac{v_{n,k}+\tilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}\right)-\rho_{\tau_k}(\tilde{\varepsilon}_{ij}-a_k)\right)。$$

其中令 $\tilde{\varepsilon}_{ij}-a_k=r$, $\dfrac{v_{n,k}+\tilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}=s$。

根据文献[13]中式(2-13)知

$$|r-s|-|r|=-s(I(r>0)-I(r<0))+2\int_0^s[I(r\leqslant t)-I(r\leqslant 0)]\mathrm{d}t,$$

则有

$$\rho_\tau(r-s)-\rho_\tau(r)=s(I(r<0)-\tau)+\int_0^s[I(r\leqslant t)-I(r\leqslant 0)]\mathrm{d}t,$$

因此,可以将 L_N 写成

$$\begin{aligned}L_N=&\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\frac{v_{n,k}+\tilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k]\\&+\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\int_0^{\frac{v_{n,k}+\tilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}}[I(\tilde{\varepsilon}_{ij}<a_k+t)-I(\tilde{\varepsilon}_{ij}<a_k)]\,\mathrm{d}t\\=&\sum_{k=1}^{K}A_N^{(k)}+\sum_{k=1}^{K}B_N^{(k)}。\end{aligned}$$

首先考虑 $B_N^{(k)}$,

$$E(B_N^{(k)})=\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\int_0^{\frac{v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}}[F(a_k+t)-F(a_k)]\,\mathrm{d}t$$

$$=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\int_0^{v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}\sqrt{n}\left[F(a_k+\frac{t}{\sqrt{n}})-F(a_k)\right]\mathrm{d}t$$

$$\to\frac{1}{2}f(a_k)(v_{n,k},\boldsymbol{\mu}_n^{\mathrm{T}})\begin{pmatrix}1&0\\0&\boldsymbol{\Sigma}\end{pmatrix}(v_{n,k},\boldsymbol{\mu}_n^{\mathrm{T}})^{\mathrm{T}},$$

$$\mathrm{Var}(B_N^{(k)})=\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}E\Big(\int_0^{\frac{v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}}(I(\tilde{\varepsilon}_{ij}\leqslant a_k+t)-I(\tilde{\varepsilon}_{ij}\leqslant a_k)-[F(a_k+t)-F(a_k)])\,\mathrm{d}t\Big)^2$$

$$\leqslant\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}E\Big(\int_0^{\frac{v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}}(I(\tilde{\varepsilon}_{ij}\leqslant a_k+t)-I(\tilde{\varepsilon}_{ij}\leqslant a_k)-[F(a_k+t)-F(a_k)])\,\mathrm{d}t\Big)\times 2\left|\frac{v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}\right|$$

$$\leqslant 4E(B_N^{(k)})\frac{\max\limits_{1\leqslant i\leqslant n,\ 1\leqslant j\leqslant l_i-q}|v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n|}{\sqrt{n}}\longrightarrow 0,$$

则

$$B_N^{(k)}\xrightarrow{p}\frac{1}{2}f(a_k)(v_{n,k},\boldsymbol{\mu}_n^{\mathrm{T}})\begin{pmatrix}1&0\\0&\boldsymbol{\Sigma}\end{pmatrix}(v_{n,k},\boldsymbol{\mu}_n^{\mathrm{T}})^{\mathrm{T}},$$

因此

$$L_N\xrightarrow{d}\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\frac{v_{n,k}+\widetilde{\boldsymbol{X}}_{ij}^{\mathrm{T}}\boldsymbol{\mu}_n}{\sqrt{n}}[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k]+\frac{1}{2}\sum_{k=1}^{K}f(a_k)(v_{n.k})^2+\frac{1}{2}\sum_{k=1}^{K}f(a_k)\boldsymbol{\mu}_n^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\mu}_n。$$

由于 L_N 是一个凸函数,因此有

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}-\boldsymbol{\beta})=\frac{\boldsymbol{\Sigma}^{-1}}{\sum_{k=1}^{K}f(a_k)}\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\frac{\widetilde{\boldsymbol{X}}_{ij}}{\sqrt{n}}[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k]+o_p(1)。$$

同时

$$E\left(\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\frac{\widetilde{\boldsymbol{X}}_{ij}}{\sqrt{n}}[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k]\right)=0,$$

$$\mathrm{Var}\left(\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{l_i-q}\frac{\widetilde{\boldsymbol{X}}_{ij}}{\sqrt{n}}\left[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k\right]\right)$$

$$=\boldsymbol{\Sigma}\,\mathrm{Var}\left(\sum_{k=1}^{K}[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k]\right)$$

$$=\boldsymbol{\Sigma}\left[\sum_{k,k'=1}^{K}\min(\tau_k,\tau_{k'})(1-\max(\tau_k,\tau_{k'}))\right]。$$

那么,由中心极限定理可得

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}-\boldsymbol{\beta})\xrightarrow{d}N\left(0,\left(\sum_{k=1}^{K}f(a_k)\right)^{-2}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_z\boldsymbol{\Sigma}^{-1}\right),$$

其中

$$\boldsymbol{\Sigma}_z=\boldsymbol{\Sigma}\mathrm{Var}\left(\sum_{k=1}^{K}[I(\tilde{\varepsilon}_{ij}<a_k)-\tau_k]\right)=\boldsymbol{\Sigma}\left[\sum_{k,k'=1}^{K}\min(\tau_k,\tau_{k'})(1-\max(\tau_k,\tau_{k'}))\right]。$$

因此

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^{\mathrm{CQR}}-\boldsymbol{\beta})\xrightarrow{d}N\left(0,\boldsymbol{\Sigma}^{-1}\frac{\sum_{k,k'=1}^{K}\min(\tau_k,\tau_{k'})(1-\max(\tau_k,\tau_{k'}))}{\left(\sum_{k=1}^{K}f(a_k)\right)^2}\right),$$

即定理得证。

2 模拟研究

本章讨论所提出方法的有限样本性质,模拟数据产生如下。

$$Y_{ij}=\boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\beta}+\boldsymbol{Z}_{ij}^{\mathrm{T}}\boldsymbol{b}_i+\varepsilon_{ij},\quad i=1,2,\cdots,n,\quad j=1,2,\cdots,l_i,$$

其中:固定效应 $\boldsymbol{\beta}=(1,2)^{\mathrm{T}}$;协变量 $\boldsymbol{X}_{ij}(i=1,2,\cdots,n,\ j=1,2,\cdots,l_i)$ 来自均值为 0、协方差矩阵为 $\begin{pmatrix}1 & 0.8\\ 0.8 & 1\end{pmatrix}$ 的多元正态分布;协变量 $Z_{ij,1}\equiv 1$, $Z_{ij,2}\sim N(0,1)(i=1,2,\cdots,n,\ j=1,2,\cdots,l_i)$;响应变量 Y_{ij} 由以上模型产生。为了探讨复合分位数回归估计的稳健性,模型误差 ε_{ij} 考虑正态分布、t 分布、柯西分布和混合正态分布这 4 种情况,即:$\varepsilon_{ij}\sim N(0,0.5^2)$、$\varepsilon_{ij}\sim 0.2t(2)$、$\varepsilon_{ij}\sim 0.2\mathrm{Cauchy}(0,1)$ 和 $\varepsilon_{ij}\sim 0.9N(0,1)+0.1N(0,10^2)$。类似文献[2],随机效应 $\boldsymbol{b}_i=(b_{i1},b_{i2})^{\mathrm{T}}$ 考虑如下情况。

(ⅰ) $\boldsymbol{b}_i\sim 0.5t(8,\boldsymbol{D})$。

(ⅱ) $b_{i1}\sim 0.5\Gamma(1,1)-0.5$, $b_{i2}\sim 0.5N(0,1)$。

(ⅲ) $\boldsymbol{b}_i\sim 0.5N(0,\boldsymbol{D})$。

$N(0,1)$ 表示标准正态分布,$N(0,\boldsymbol{D})$ 是均值为 0、协方差矩阵为 $\boldsymbol{D}$ 的正态分布,$\Gamma(1,1)$ 为尺度和形状参数均为 1 的 Gamma 分布,$t(8)$ 为 8 个自由度的 t 分布,$t(8,D)$ 为自由度为 8、协方差矩阵为 $\boldsymbol{D}$ 的 t 分布,Cauchy(0,1)是位置参数为 0、尺度参数为 1 的柯西分布,$\boldsymbol{D}=\begin{pmatrix}1 & 0.5\\ 0.5 & 1\end{pmatrix}$。样本容量分别取 $n=100,200,300$,每个个体重复观测次数 $l_i=5$,对每种情况重复运行 500 次。

为了比较,考虑了 2 种方法:Wu 等[2]提出的基于估计方程的正交矩估计方法(orthogonality-based estimation of moment, OBE)和本文提出的基于正交投影的复合分位数回归估计方法(composite quantile regression estimation based on orthogonal projection, CQRO),分位点个数分别取 $K=5,9$,相应估计记为 CQRO-5和 CQRO-9。为了评估估计的精度,计算了标准差(standard deviation,SD)和绝对偏差(absolute deviation,AD),模拟结果见表 1—4。表 1—4 给出了当 ε_{ij} 和 $\boldsymbol{b}_i$ 分别服从不同分布时,采用 OBE、CQRO-5 和 CQRO-9 这 3 种方法对固定效应估计的结果。

表 1 基于不同的估计方法得到的对固定效应估计的标准差和绝对偏差($\varepsilon_{ij}\sim N(0,0.5^2)$)

Table 1 Standard deviation and absolute deviation of fixed effect estimation based on different estimation methods ($\varepsilon_{ij}\sim N(0,0.5^2)$)

n	随机效应服从的模型	方法	β_1		β_2	
			SD	AD	SD	AD
100	(ⅰ)	CQRO-5	0.045 9	0.036 8	0.047 7	0.038 0
		CQRO-9	0.045 1	0.036 2	0.046 9	0.037 2
		OBE	0.044 8	0.035 9	0.046 5	0.037 2
	(ⅱ)	CQRO-5	0.049 7	0.040 2	0.0491	0.039 8
		CQRO- 9	0.0491	0.039 9	0.048 2	0.039 0
		OBE	0.0474	0.038 3	0.047 0	0.0382
	(ⅲ)	CQRO-5	0.049 5	0.040 0	0.050 9	0.041 4
		CQRO- 9	0.049 2	0.039 5	0.050 1	0.040 6
		OBE	0.048 8	0.039 2	0.050 4	0.040 7
200	(ⅰ)	CQRO-5	0.034 6	0.027 5	0.033 8	0.026 6
		CQRO- 9	0.0343	0.027 2	0.033 6	0.026 2
		OBE	0.034 0	0.026 8	0.032 9	0.025 6
	(ⅱ)	CQRO-5	0.033 7	0.026 6	0.033 6	0.027 0
		CQRO- 9	0.033 5	0.026 6	0.033 2	0.0267
		OBE	0.033 2	0.026 4	0.032 5	0.026 2
	(ⅲ)	CQRO-5	0.0344	0.0266	0.036 0	0.029 0
		CQRO- 9	0.0344	0.026 8	0.035 4	0.028 5
		OBE	0.033 0	0.025 7	0.034 3	0.027 7

表 1(续)

n	随机效应服从的模型	方法	β_1		β_2	
			SD	AD	SD	AD
300	(ⅰ)	CQRO-5	0.027 5	0.021 9	0.028 0	0.021 9
		CQRO- 9	0.026 8	0.0213	0.027 4	0.021 5
		OBE	0.0264	0.020 9	0.026 7	0.020 9
	(ⅱ)	CQRO-5	0.029 3	0.023 6	0.027 6	0.021 9
		CQRO-9	0.028 7	0.023 0	0.027 2	0.0215
		OBE	0.028 5	0.023 0	0.027 0	0.021 2
	(ⅲ)	CQRO-5	0.028 0	0.022 8	0.027 6	0.022 2
		CQRO-9	0.028 0	0.022 7	0.0276	0.022 0
		OBE	0.027 4	0.022 1	0.027 3	0.021 8

表 2　基于不同的估计方法得到的对固定效应估计的标准差和绝对偏差($\varepsilon_{ij}\sim 0.2t(2)$)

Table 2　Standard deviation and absolute deviation of fixed effect estimation based on different estimation methods ($\varepsilon_{ij}\sim 0.2t(2)$)

n	随机效应服从的模型	方法	β_1		β_2	
			SD	AD	SD	AD
100	(ⅰ)	CQRO-5	0.034 9	0.028 2	0.034 1	0.026 9
		CQRO-9	0.035 0	0.028 1	0.033 9	0.026 9
		OBE	0.065 5	0.046 1	0.061 4	0.044 1
	(ⅱ)	CQRO-5	0.037 5	0.029 5	0.0366	0.0296
		CQRO-9	0.037 2	0.0296	0.036 8	0.0297
		OBE	0.065 8	0.0453	0.067 0	0.049 0
	(ⅲ)	CQRO-5	0.034 2	0.027 5	0.034 1	0.027 2
		CQRO-9	0.033 9	0.0276	0.034 5	0.027 4
		OBE	0.067 5	0.045 5	0.064 0	0.046 2
200	(ⅰ)	CQRO-5	0.025 4	0.020 0	0.025 0	0.019 9
		CQRO-9	0.025 4	0.020 2	0.025 2	0.020 1
		OBE	0.0416	0.031 0	0.043 8	0.033 0
	(ⅱ)	CQRO-5	0.024 8	0.020 0	0.024 6	0.019 8
		CQRO-9	0.025 1	0.020 2	0.024 9	0.019 9
		OBE	0.068 5	0.037 3	0.073 3	0.037 9
	(ⅲ)	CQRO-5	0.024 9	0.019 7	0.025 0	0.019 9
		CQRO-9	0.024 9	0.019 9	0.025 0	0.019 9
		OBE	0.0444	0.033 5	0.045 8	0.033 9
300	(ⅰ)	CQRO-5	0.020 1	0.015 7	0.021 0	0.016 5
		CQRO-9	0.020 5	0.016 0	0.021 1	0.016 4
		OBE	0.042 9	0.029 0	0.045 8	0.029 7
	(ⅱ)	CQRO-5	0.021 5	0.017 1	0.0206	0.016 4
		CQRO-9	0.0214	0.017 0	0.020 5	0.016 4
		OBE	0.039 5	0.0294	0.0394	0.028 6
	(ⅲ)	CQRO-5	0.020 4	0.016 4	0.020 3	0.016 2
		CQRO-9	0.020 3	0.016 3	0.020 3	0.016 2
		OBE	0.039 8	0.028 9	0.037 0	0.028 1

从表 1—4 可以看出:1) 当模型误差服从正态分布时,OBE 和 CQRO 的模拟结果很接近。当模型误差服从非正态分布时,CQRO 的模拟结果要明显优于 OBE 的,说明 CQRO 估计方法比 OBE 估计方法更稳健。

2) 当复合分位数回归估计方法中的分位点个数 $K=5$, 9 时,无论模型误差服从什么分布,CQRO-5 和 CQRO-9 两者的结果都是类似的,这也验证了 Zou 等[6]指出的当 $K\geqslant 5$ 时的估计结果都比较好且不受 K 的取值的影响的结论,因此在实际应用中只须要选择 $K\geqslant 5$ 的值即可。

表 3 基于不同的估计方法得到的对固定效应估计的标准差和绝对偏差($\varepsilon_{ij}\sim 0.2\text{Cauchy}(0,1)$)

Table 3 Standard deviation and absolute deviation of fixed effect estimation based on different estimation methods ($\varepsilon_{ij}\sim 0.2\text{Cauchy}(0,1)$)

n	随机效应服从的模型	方法	β_1		β_2	
			SD	AD	SD	AD
100	(ⅰ)	CQRO-5	0.071 0	0.055 2	0.067 6	0.054 4
		CQRO-9	0.071 2	0.055 5	0.068 8	0.055 3
		OBE	9.031 7	1.201 9	8.025 0	1.084 8
	(ⅱ)	CQRO-5	0.070 2	0.055 9	0.072 7	0.058 9
		CQRO-9	0.069 8	0.055 6	0.072 2	0.058 5
		OBE	8.139 6	1.414 8	10.662 0	1.781 0
	(ⅲ)	CQRO-5	0.070 9	0.057 2	0.069 9	0.055 1
		CQRO-9	0.071 8	0.058 1	0.070 3	0.055 4
		OBE	4.598 1	1.306 7	4.098 2	1.216 5
200	(ⅰ)	CQRO-5	0.051 6	0.041 6	0.050 5	0.040 2
		CQRO-9	0.052 7	0.042 6	0.051 2	0.041 1
		OBE	89.829 0	5.894 3	45.244 0	3.891 6
	(ⅱ)	CQRO-5	0.051 3	0.041 0	0.050 9	0.040 6
		CQRO-9	0.052 0	0.041 9	0.051 8	0.041 8
		OBE	4.716 6	1.289 6	4.461 4	1.205 6
	(ⅲ)	CQRO-5	0.048 9	0.038 7	0.049 2	0.038 6
		CQRO-9	0.0493	0.038 8	0.498 0	0.039 1
		OBE	9.347 7	1.588 2	32.851 7	2.923 4
300	(ⅰ)	CQRO-5	0.041 6	0.033 6	0.041 1	0.033 7
		CQRO-9	0.041 6	0.0333	0.040 6	0.033 2
		OBE	10.131 7	1.548 9	4.471 8	1.250 0
	(ⅱ)	CQRO-5	0.038 2	0.029 8	0.039 2	0.031 3
		CQRO-9	0.039 0	0.030 7	0.039 2	0.031 6
		OBE	11.6173	1.694 9	8.124 6	1.561 0
	(ⅲ)	CQRO-5	0.040 2	0.031 1	0.0413	0.033 2
		CQRO-9	0.040 7	0.031 5	0.041 8	0.033 6
		OBE	14.961 7	2.088 7	10.626 3	1.888 7

表 4 基于不同的估计方法得到的对固定效应估计的标准差和绝对偏差($\varepsilon_{ij}\sim 0.9N(0,1)+0.1N(0,10^2)$)

Table 4 Standard deviation and absolute deviation of fixed effect estimation based on different estimation methods ($\varepsilon_{ij}\sim 0.9N(0,1)+0.1N(0,10^2)$)

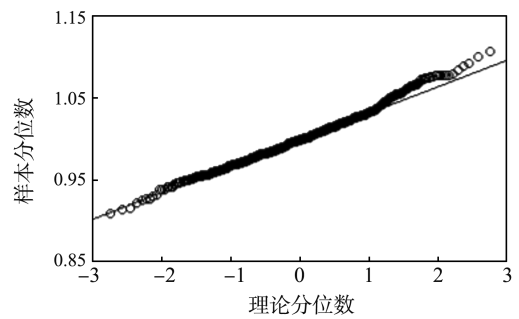
n	随机效应服从的模型	方法	β_1		β_2	
			SD	AD	SD	AD
100	(ⅰ)	CQRO-5	0.118 5	0.094 5	0.113 8	0.091 6
		CQRO-9	0.115 5	0.092 4	0.112 1	0.089 7
		OBE	0.299 1	0.238 0	0.294 6	0.234 2
	(ⅱ)	CQRO-5	0.125 5	0.100 5	0.117 1	0.092 4
		CQRO-9	0.123 8	0.098 5	0.117 3	0.092 0
		OBE	0.320 1	0.258 1	0.315 1	0.251 4
	(ⅲ)	CQRO-5	0.123 1	0.097 4	0.127 8	0.100 0
		CQRO-9	0.1223	0.096 7	0.127 5	0.099 6
		OBE	0.335 8	0.269 1	0.336 0	0.265 5
200	(ⅰ)	CQRO-5	0.083 6	0.067 9	0.082 1	0.065 5
		CQRO-9	0.0833	0.067 8	0.082 7	0.066 3
		OBE	0.222 0	0.177 2	0.217 2	0.172 4
	(ⅱ)	CQRO-5	0.083 9	0.068 5	0.088 7	0.072 1
		CQRO-9	0.085 4	0.069 0	0.089 6	0.073 0
		OBE	0.230 7	0.185 9	0.234 6	0.188 6
	(ⅲ)	CQRO-5	0.0833	0.067 2	0.082 1	0.066 5
		CQRO-9	0.082 9	0.067 2	0.081 9	0.066 7
		OBE	0.217 2	0.171 4	0.219 0	0.173 6

表 4(续)

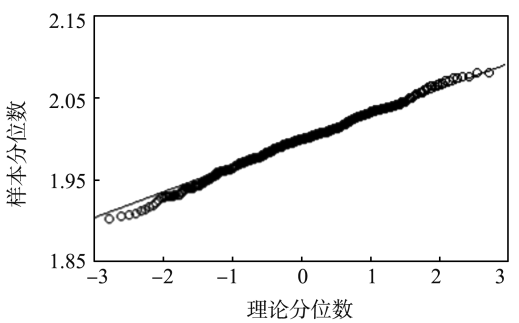
n	随机效应服从的模型	方法	β_1		β_2	
			SD	AD	SD	AD
300	(ⅰ)	CQRO- 5	0.072 3	0.058 5	0.074 2	0.059 8
		CQRO-9	0.0723	0.058 4	0.073 7	0.059 7
		OBE	0.186 2	0.151 8	0.194 8	0.157 0
	(ⅱ)	CQRO-5	0.068 8	0.055 1	0.066 7	0.053 9
		CQRO-9	0.068 9	0.054 6	0.066 8	0.053 3
		OBE	0.1913	0.153 5	0.185 4	0.150 1
	(ⅲ)	CQRO-5	0.067 0	0.053 6	0.067 1	0.052 7
		CQRO-9	0.066 3	0.053 2	0.067 6	0.053 1
		OBE	0.185 4	0.149 3	0.186 2	0.145 5

3) 随着样本容量的增加,CQRO 得到的估计值的 SD 和 AD 都在减小。

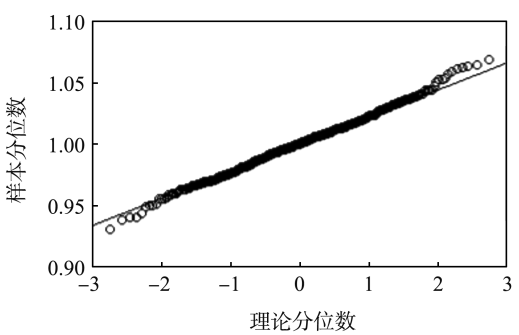
其次,为了验证本文提出的 CQRO 估计是否具有渐近正态性,绘制了 β_1 和 β_2 重复运行 500 次所得的 CQRO-5 和 CQRO-9 估计的 Q-Q 图。图 1、2 仅给出了样本量 $n=200$ 时 4 种不同模型误差分布情形的Q-Q 图。对样本量 $n=100,300$ 结果类似。从图 1、2 给出的 500 次模拟出的 $\boldsymbol{\beta}$ 的 CQRO-5 和 CQRO-9 估计 Q-Q 图可以看出,Q-Q 图上的点近似地在一条直线附近,因此所提出的 CQRO-5 和 CQRO-9 估计具有渐近正态性。



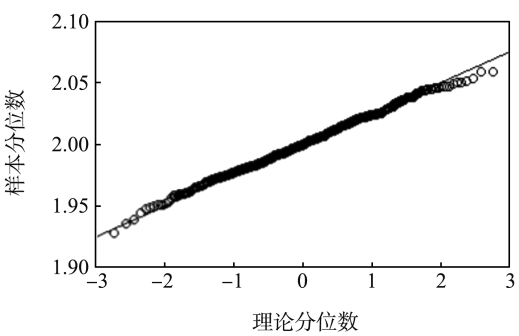
(a) $\varepsilon_{ij}\sim N(0,0.5^2)(\beta_1)$



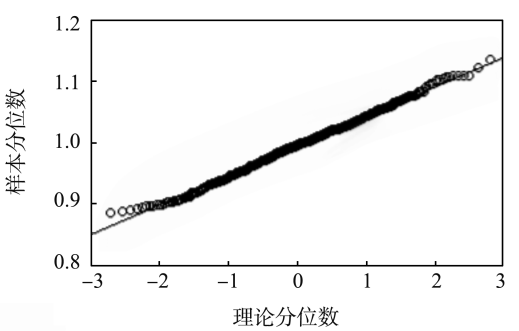
(b) $\varepsilon_{ij}\sim N(0,0.5^2)(\beta_2)$



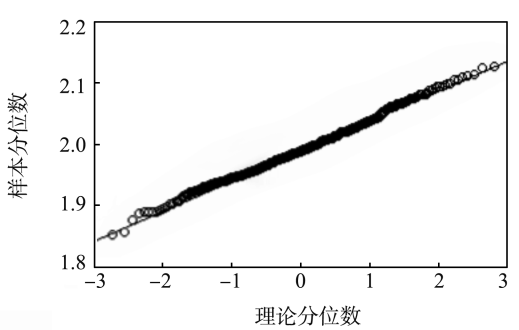
(c) $\varepsilon_{ij}\sim 0.2t(2)(\beta_1)$



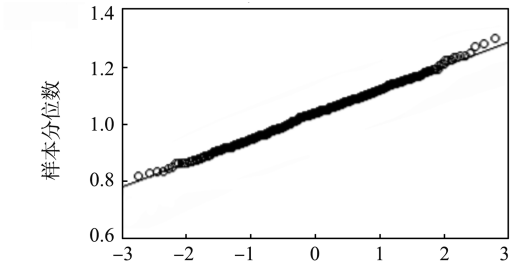
(d) $\varepsilon_{ij}\sim 0.2t(2)(\beta_2)$



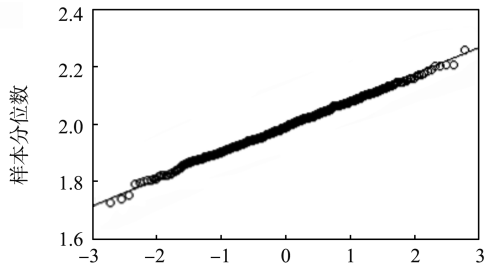
(e) $\varepsilon_{ij}\sim 0.2\text{Cauchy}(0,1)(\beta_1)$



(f) $\varepsilon_{ij}\sim 0.2\text{Cauchy}(0,1)(\beta_2)$



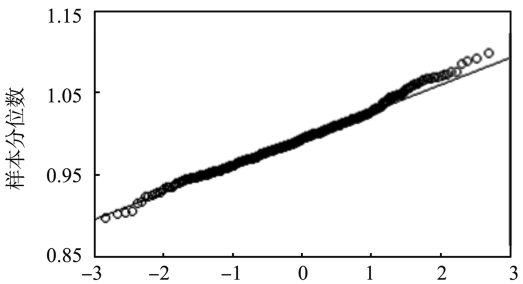
(g) ε_{ij}~0.9N(0,1)+0.1N(0,10^2)(β_1)



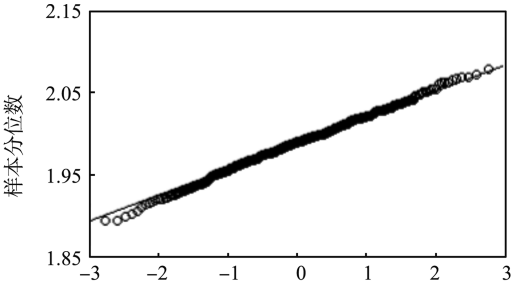
(h) ε_{ij}~0.9N(0,1)+0.1N(0,10^2)(β_2)

图 1 样本容量 $n=200$ 时 4 种模型误差分布情形下 CQRO-5 估计的 Q-Q 图

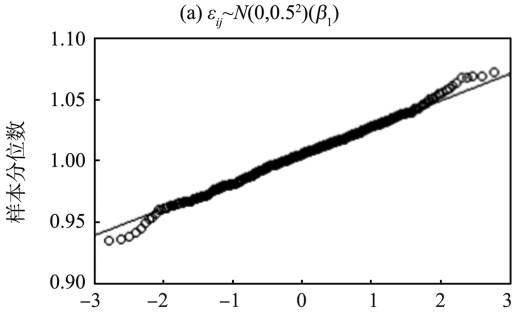
Fig.1 Q-Q plot of CQRO-5 estimation with four model error distributions when sample size $n=200$



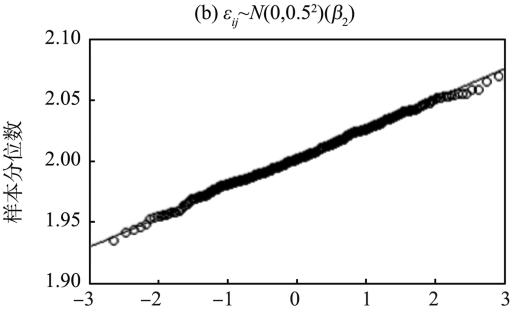
(a) ε_{ij}~N(0,0.5^2)(β_1)



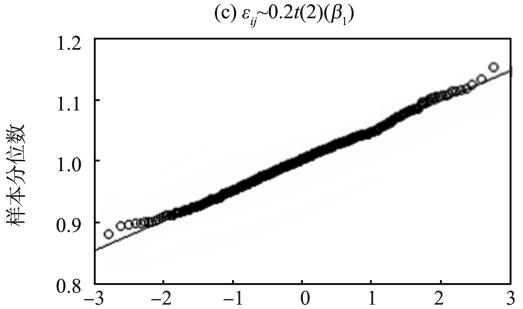
(b) ε_{ij}~N(0,0.5^2)(β_2)



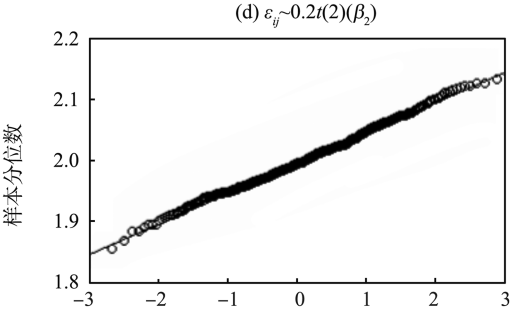
(c) ε_{ij}~0.2t(2)(β_1)



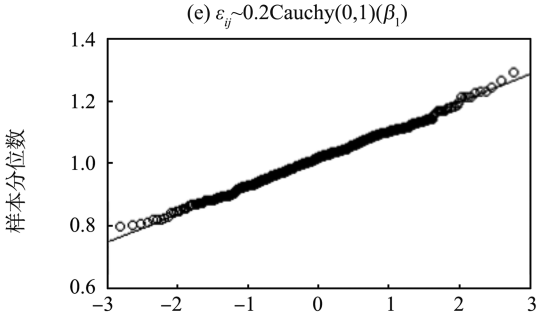
(d) ε_{ij}~0.2t(2)(β_2)



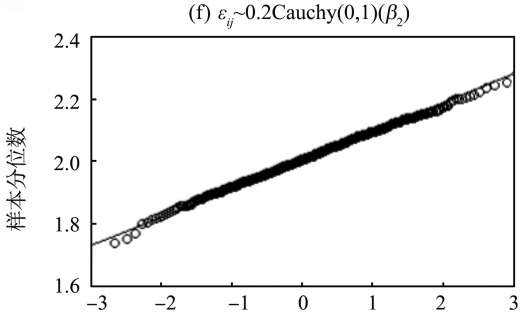
(e) ε_{ij}~0.2Cauchy(0,1)(β_1)



(f) ε_{ij}~0.2Cauchy(0,1)(β_2)



(g) ε_{ij}~0.9N(0,1)+0.1N(0,10^2)(β_1)



(h) ε_{ij}~0.9N(0,1)+0.1N(0,10^2)(β_2)

图 2 样本量 $n=200$ 时 4 种模型误差分布情形下 CQRO-9 估计的 Q-Q 图

Fig.2 Q-Q plot of CQRO-9 estimation with four model error distributions when sample size $n=200$

最后,为了比较 OBE、CQRO-5 和 CQRO-9 估计的性能,基于样本容量 $n=200$,绘制了 β_1 和 β_2 重复运行 500 次所得的 OBE、CQRO-5 和 CQRO-9 估计的 L_1 估计误差和 L_2 估计误差的箱线图,其中 L_1 估计误差和 L_2 估计误差的定义为 $L_1:|\hat{\beta}_1-\beta_{01}|+|\hat{\beta}_2-\beta_{02}|$,$L_2:|\hat{\beta}_1-\beta_{01}|^2+|\hat{\beta}_2-\beta_{02}|^2$。对样本容量 $n=100$,300 结果类似。从图 3、4 可以看出,在模型误差服从正态分布时,OBE、CQRO-5 和 CQRO-9 估计的 L_1 估计误差和 L_2 估计误差相差不大;在模型误差服从非正态分布时,CQRO 明显优于 OBE。

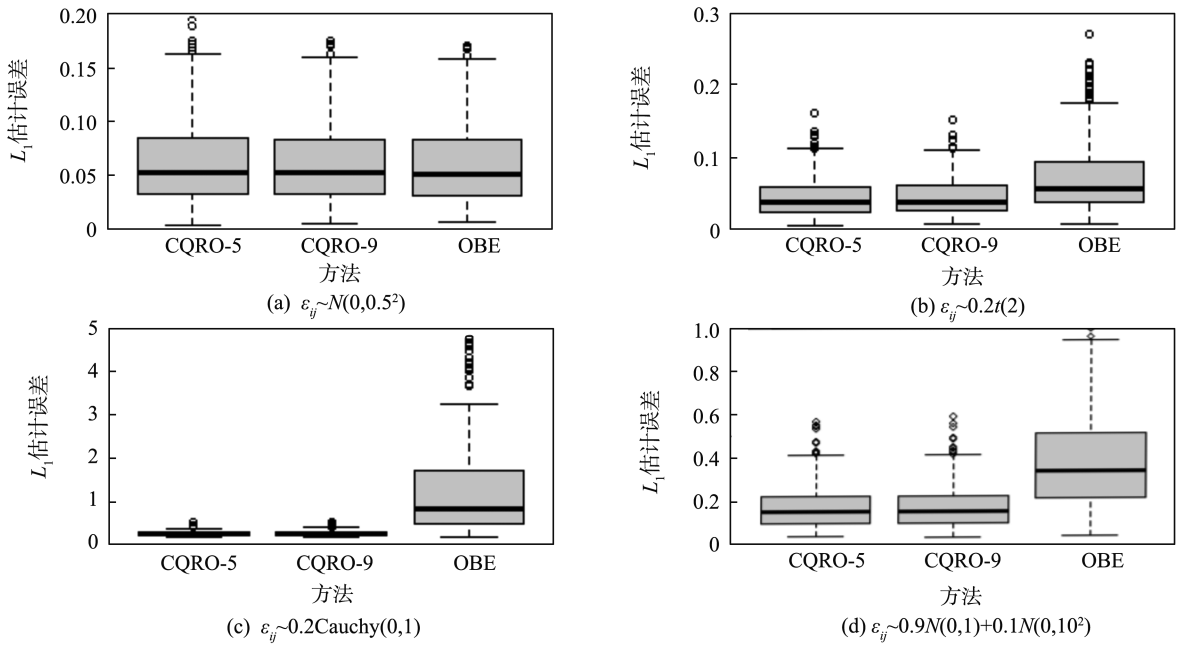


图 3　样本容量 $n=200$ 时 4 种模型误差分布情形下 CQRO-5、CQRO-9 和 OBE 估计的 L_1 估计误差箱线图

Fig.3　Boxplot of L_1 estimation errors of CQRO-5, CQRO-9 and OBE estimators with four model error distributions when the sample size $n=200$

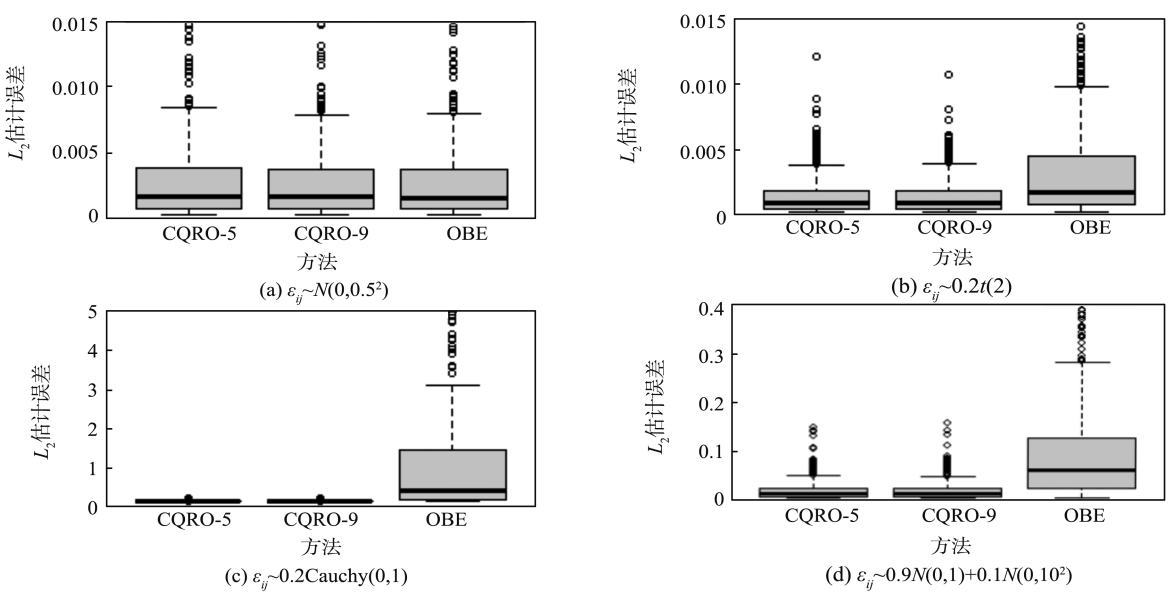


图 4　样本容量 $n=200$ 时 4 种模型误差分布情形下 CQRO-5、CQRO-9 和 OBE 估计的 L_2 估计误差箱线图

Fig.4　Boxplot of L_2 estimation errors of CQRO-5, CQRO-9 and OBE estimators with four model error distributions when the sample size $n=200$

3　实例分析

本章分析了 R 软件包 plm 中的 Produce 数据集。该数据集收集了 1970—1986 年美国 48 个州的经济增长相关指标,其中“hwy”表示高速公路和街道,“water”表示供水和排污水系统,“pc”表示个人资本存

量,“gsp”表示州生产总值,“emp”表示劳动力投入及“umemp”表示州失业率。Munnell[14]利用该数据集分析了公共资本对经济增长的影响,分别把高速公路和街道、供水和排污水系统视为公共资本进行了分析。本文也考虑公共资本对该经济增长的影响,建模如下线性混合效应模型:

$$y_{\mathrm{gsp}_{ij}}=x_{\mathrm{water}_{ij}}\beta_1+x_{\mathrm{hwy}_{ij}}\beta_2+b_i+\varepsilon_{ij},\quad i=1,2,\cdots,48,\quad j=1,2,\cdots,17,$$

其中 y_{gsp}, x_{water}, x_{hwy}进行了对数变换。

类似模拟研究,采用 OBE、CQRO 方法分析了该数据集,计算结果见表 5。从表 5 可以看出,CQRO-5 和 CQRO-9 的估计结果类似,与 OBE 方法相比,高速公路对经济增长的正影响更强一些。图 5 给出了州生产总值的直方图和密度图,从图 5 可以看出,州生产总值的分布并非正态分布。因此,与 OBE 方法相比,用 CQRO 方法分析该数据可能更合理,因为 CQRO 方法具有稳健性。

表 5 对 Produce 数据集 OBE、CQRO-5 和 CQRO-9 方法回归系数的估计

Table 5 Estimators of regression coefficients for OBE, CQRO-5 and CQRO-9 methods in the Produce dataset

方法	OBE	CQRO-5	CQRO-9
β_1	0.523 7	0.516 6	0.510 8
β_2	0.534 0	0.541 8	0.545 1

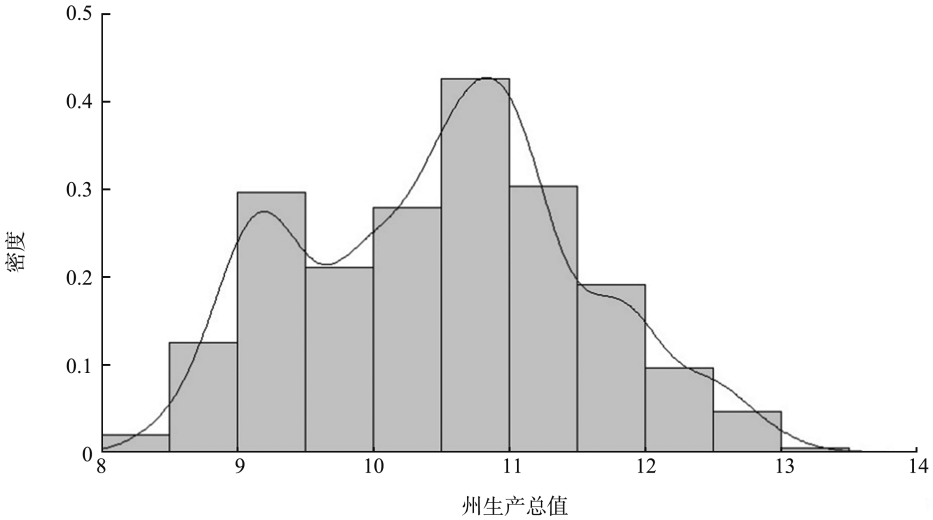


图 5 州生产总值的直方图和密度图

Fig.5 Histograms and density graphs of the gross state product

4 结论

线性混合效应模型是分析纵向数据常见的模型之一。本文针对线性混合效应模型提出了一种基于 QR 分解的复合分位数回归估计。在一些正则条件下,给出了固定效应估计的渐近性质,模拟研究了本文提出方法的有限样本性质,并将其应用于实际数据分析。本文提出的方法具有 2 个优势:(1) 巧妙地通过 QR 分解技术消除随机效应,使得固定效应和随机效应分离,彼此互不影响;(2) 利用复合分位数回归估计方法给出固定效应的估计,与正交矩估计方法相比,本文提出的方法更稳健。需要注意的是,本文提出的复合分位数回归估计的渐近方差含有未知的 $f(\cdot)$,如何去获得渐近方差的估计是值得进一步讨论的。除此之外,本文仅考虑了固定效应的估计,如何结合 QR 分解技术和复合分位数回归来研究高维混合效应模型的变量选择也是将要考虑的问题。

参考文献:

[1] CUI H J, NG K W, ZHU L X. Estimation in mixed effects model with errors in variables[J]. Journal of Multivariate Analysis, 2004, 91(1):53-73.

[2] WU Ping, ZHU Lixing. An orthogonality-based estimation of moments for linear mixed models[J]. Scandinavian Journal of

Statistics,2010,37:253-263.
[3] 陈心洁,林鹏,邹国华. 线性混合效应模型的 FIC 选择准则[J]. 统计研究,2015,32(3):100-103.
CHEN Xinjie, LIN Peng, ZOU Guohua. FIC selection criteria for linear mixed effects models[J]. Statistical Research, 2015, 32(3):100-103.
[4] 林鹏. 一般线性混合效应模型的随机效应选择研究[J]. 系统科学与数学,2015,35(6):617-626.
LIN Peng. Research on random effects selection for general linear mixed effects model[J]. Journal of Systems Science and Mathematical Sciences, 2015, 35(6):617-626.
[5] 赵培信,张帆,周小双. 不完全观测数据下混合效应模型的正交投影估计[J]. 工程数学学报,2023,40(1):97-109.
ZHAO Peixin, ZHANG Fan, ZHOU Xiaoshuang. Orthogonal projection estimation for mixed effects models with incomplete observations data [J]. Chinese Journal of Engineering Mathematics, 2023, 40(1):97-109.
[6] ZOU Hui, YUAN Ming. Composite quantile regression and the oracle model selection theory[J]. The Annals of Statistics, 2008, 36(3):1108-1126.
[7] 王康宁,李劭珉,林路. 基于 copula 函数的纵向数据复合分位数回归及变量选择[J]. 中国科学(数学),2020,50(8):1097-1116.
WANG Kangning, LI Shaomin, LIN Lu. Composite quantile regression and variable selection of longitudinal data based on copula function [J]. Science China Mathematics, 2020, 50(8):1097-1116.
[8] 刘艳霞,芮荣祥,田茂再. 部分线性变系数模型的新复合分位数回归估计[J]. 应用数学学报,2021,44(2):159-174.
LIU Yanxia, RUI Rongxiang, TIAN Maozai. New composite quantile regression estimation for partial linear variable coefficient models [J]. Acta Mathematicae Applicatae Sinica, 2021, 44(2):159-174.
[9] 张永霞,田茂再. 基于贝叶斯的部分线性单指标复合分位回归的研究及其应用[J]. 系统科学与数学,2021,41(5):1381-1399.
ZHANG Yongxia, TIAN Maozai. Research and application of partial linear single index composite quantile regression based on Bayes [J]. Journal of Systems Science and Mathematical Sciences, 2021, 41(5):1381-1399.
[10] 张立文,程东坡,薛文骏,等. 复合分位数下门限自回归模型的变点估计[J]. 中国科学(数学),2022,52(1):63-84.
ZHANG Liwen, CHENG Dongpo, XUE Wenjun, et al. Change point estimation of autoregressive model with lower threshold of composite quantile [J]. Science China Mathematics, 2022, 52(1): 63-84.
[11] JIANG Rong, SUN Mengxian. Single-index composite quantile regression for ultra-high-dimensional date[J]. TEST, 2022, 31(2):443-460.
[12] GUO Chaohui, LYU Jing, WU Jibo. Composite quantile regression for ultra-high dimensional semiparametric modle averaging[J]. Computational Statistics & Data Analysis, 2021, 160:107231.
[13] KNIGHT K. Limiting distributions for l_1 regression estimators under general conditions[J]. The Annals of Statistics, 1998, 26(2):755-770.
[14] MUNNELL A. Why has productivity growth declined productivity and public investment[J]. New England Economic Review, 1990, 1(2):3-22.

(编辑:李艺)